Kannan Ramchandran[*]
Rohit Puri[*]
Jiajun Wang[*]
Abhik Majumdar[*]
Marco Tagliasacchi[+]
Stefano Tubaro[+]

* University of California, Berkeley, USA,
+ Politecnico di Milano, Italy

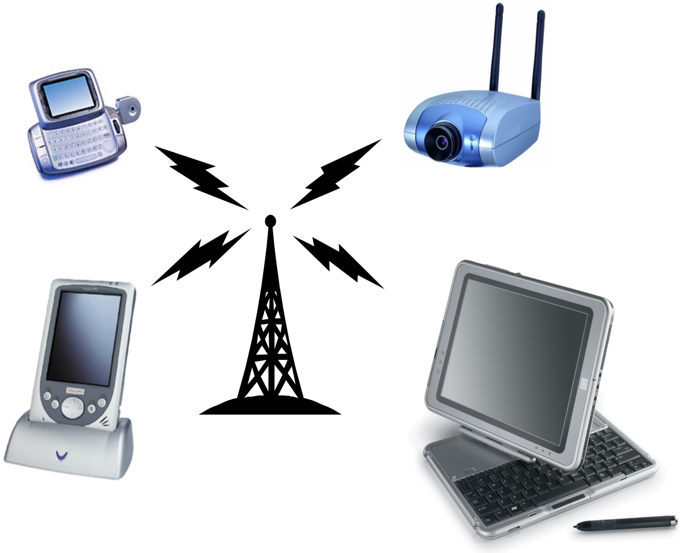
Today’s video codec architectures

- Driven primarily by downlink (broadcast) model:
  - Complex transmitter / Light receiver
  - Not explicitly designed for lossy channels

The changing landscape…

New class of “uplink” media applications:

- High-resolution wireless digital video cameras
- Multimedia-mobile phones & PDA’s
- Ultra-low-power video sensors and surveillance cameras
- Wireless-video teleconferencing systems
- Home-entertainment and home-networking systems



Challenges

- Robustness to packet/frame
- Light codec complexity
- Low delay
- High compression efficiency

Meet these requirements *simultaneously*

What about today’s error-resiliency toolkit?

- Available toolkit:
  - Decoder error concealment
  - Robust entropy coding: RVLC,...
  - Error-resilient prediction: intra-refresh,...
  - Layering with Unequal Error Protection
  - R-D optimization

  ...

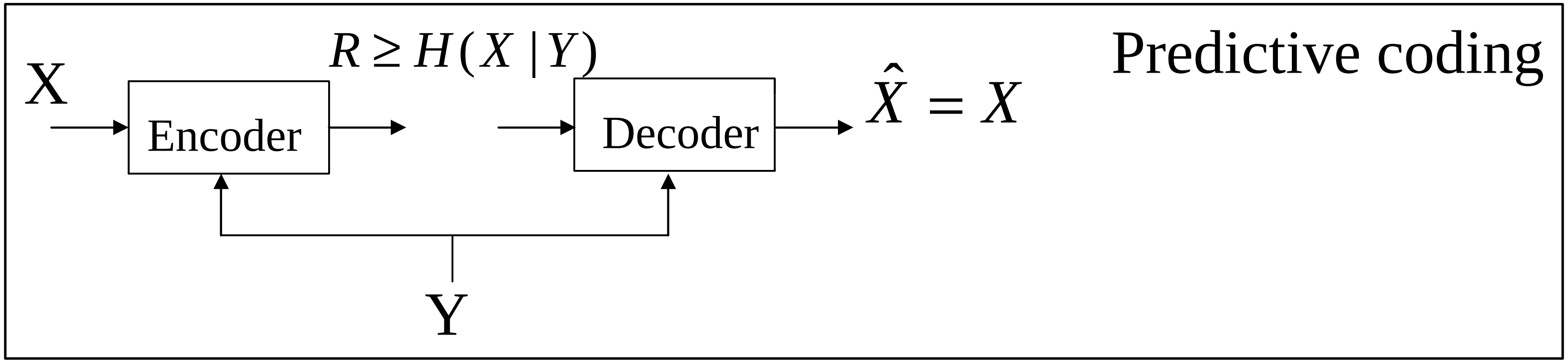
- Useful but **root cause** of problems not addressed:
  - **Predictive motion-compensation framework is fundamentally fragile**

Go from deterministic to statistical mindset

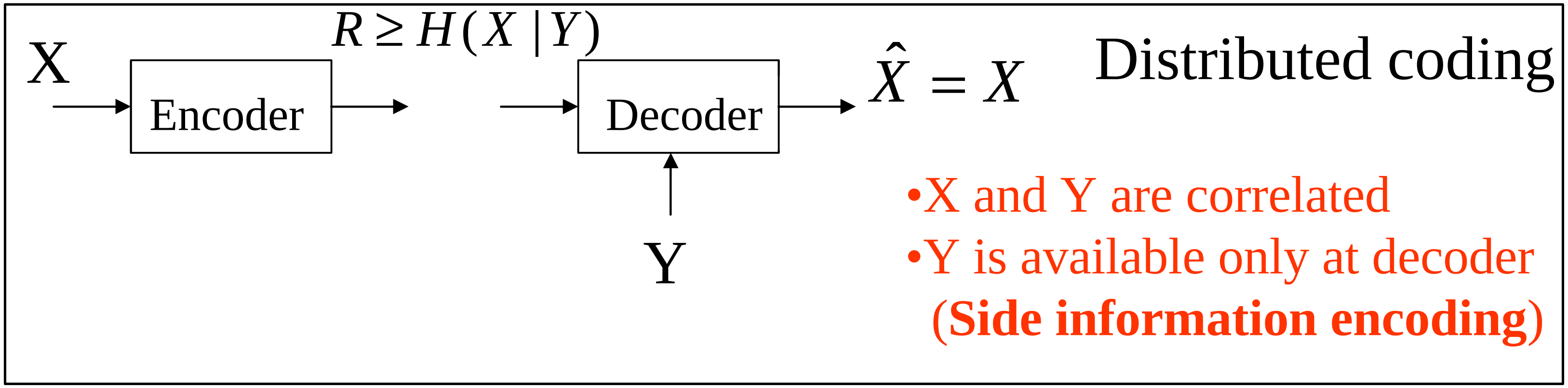
Is there anything missing in our toolkit?

Distributed Source Coding

Distributed source coding



Predictive coding



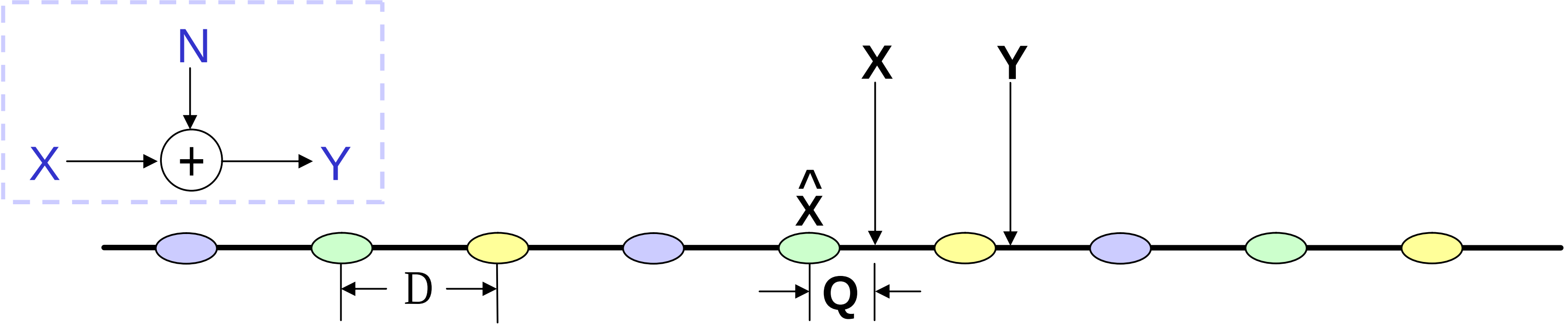
Distributed coding

•X and Y are correlated

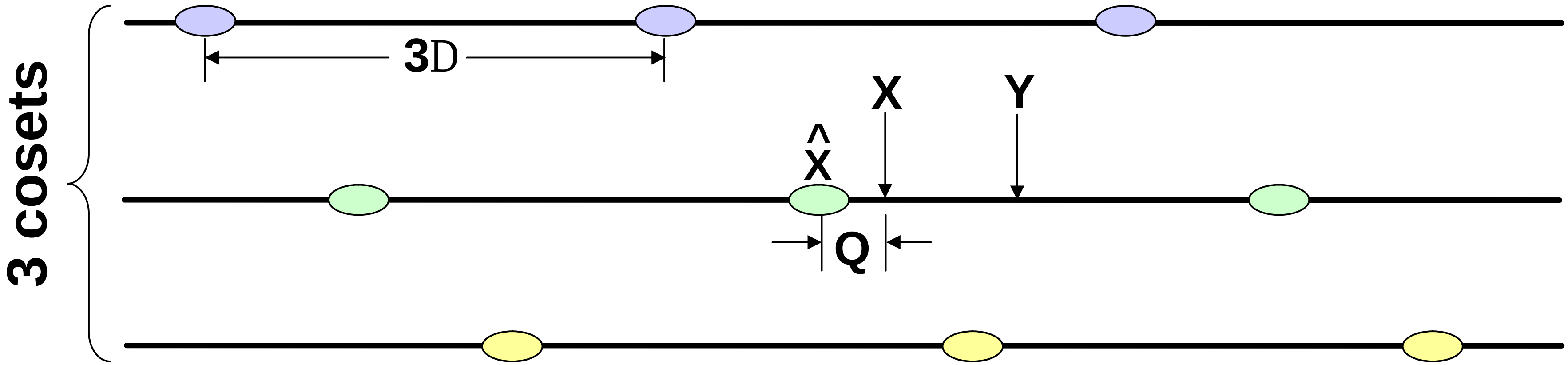
•Y is available only at decoder
(Side information encoding)

- Slepian&Wolf ('73) (lossless)
- Wyner&Ziv ('76) (lossy)

Wyner-Ziv framework



Partition



- Encoder: send the index of the coset ($\log_2 3$ bits)
- Decoder: decode X based on Y and signaled coset

Application Landscape For SI Coding

- Sensor networks
- Digital upgrade of legacy analog systems
- M-channel Multiple Description codes
- Media broadcast using hybrid analog/digital techniques
- Multimedia security: Data-hiding, watermarking, steganography
- Compression of encrypted data
- Video Coding

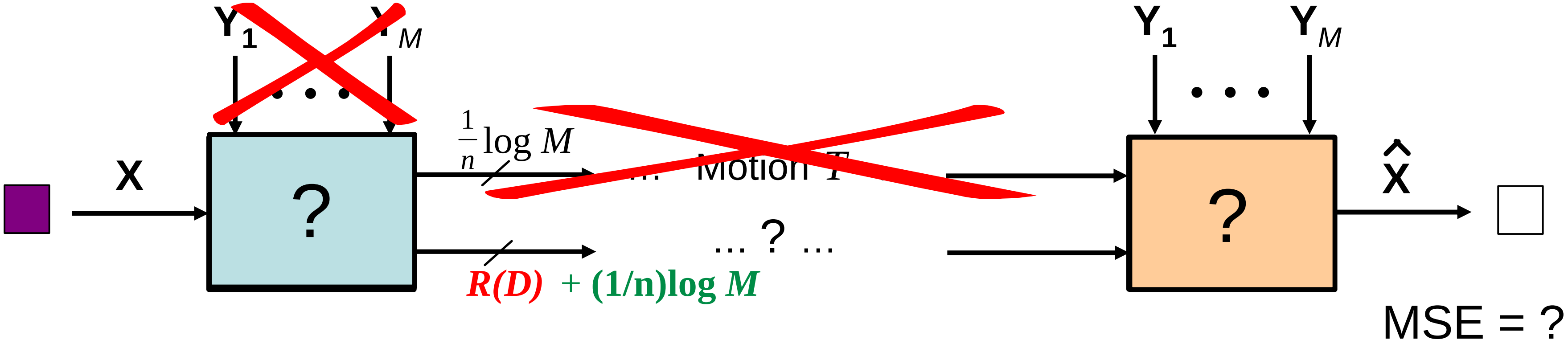
From information theory to video coding

Distributed video coding architecture: **PRISM**

(**P**ower-efficient, **R**obust, h**I**gh compression **S**yndrome based **M**ultimedia coding)

(Puri & Ramchandran: Allerton '02)

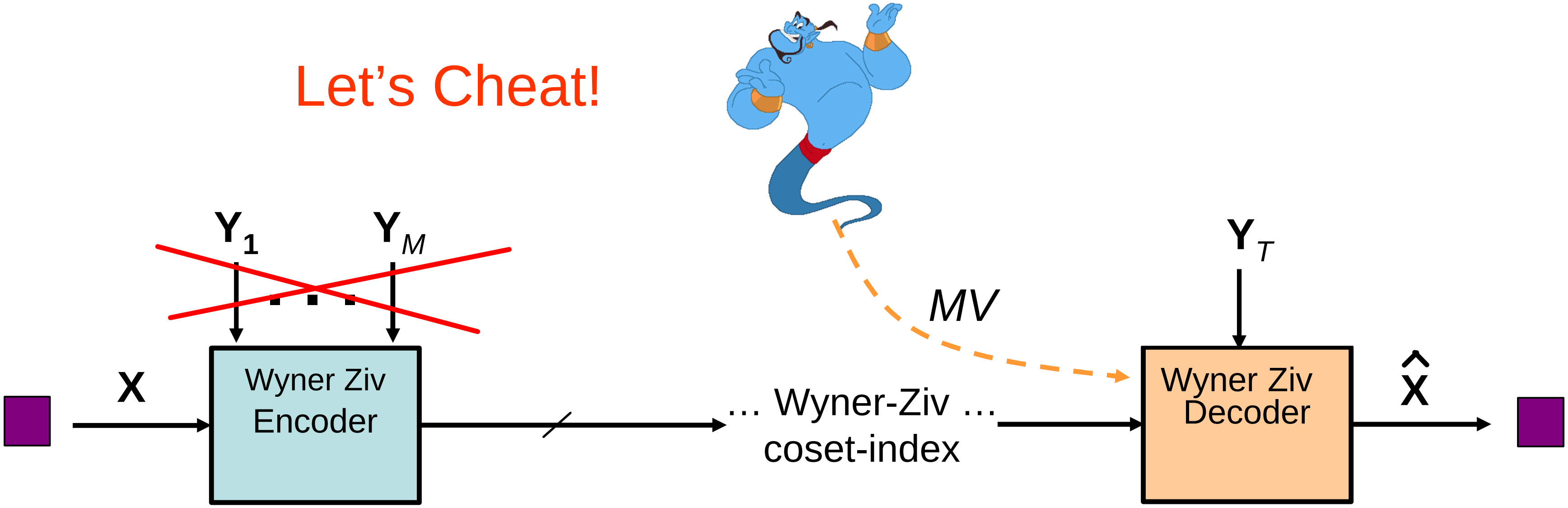
Motion-free encoding?



- The encoder does not have or cannot use $\mathbf{Y}_1, \ldots, \mathbf{Y}_M$ and
- The decoder does not know *T.*
- The encoder may work at rate: ***R(D)*** + ***(1/n)*log *M*** bits per pixel.
- How to decode and what is the performance?

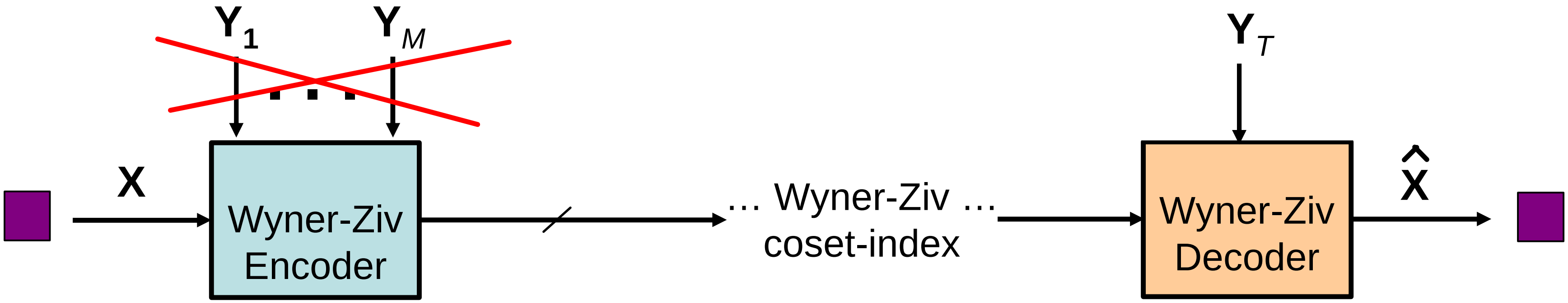
Is a No-Motion Encoder Possible?

Let's Cheat!



- Let's cheat & let the decoder have the MV → "classical" W-Z problem
- The encoder works at same rate as predictive coder

Is a No-Motion Encoder Possible?



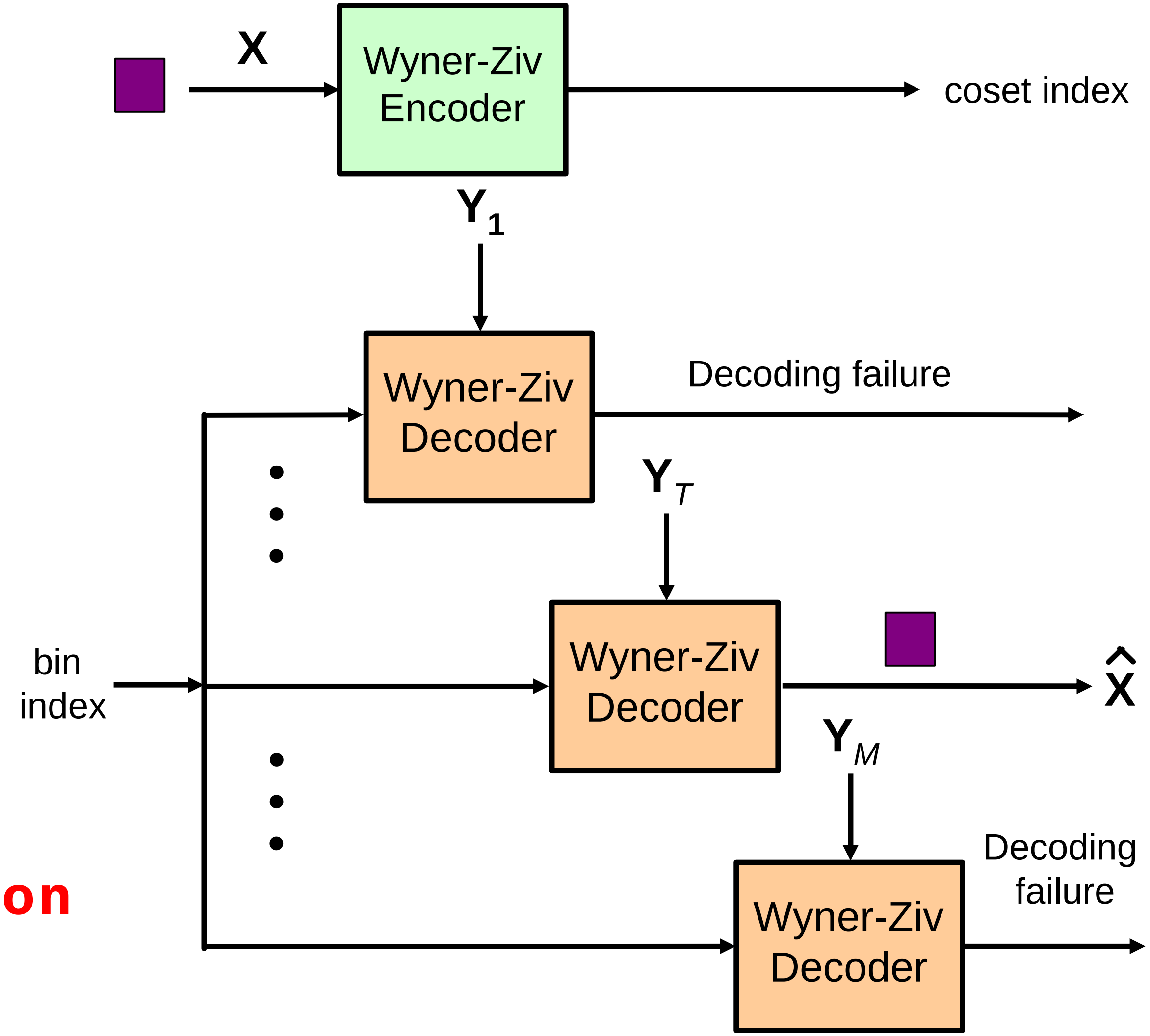
- Can decoding work without a genie?
  - *Yes*
- Can we match the performance of predictive coding?
  - *Yes* (when DFD statistics are Gaussian)

No Genie, No Problem

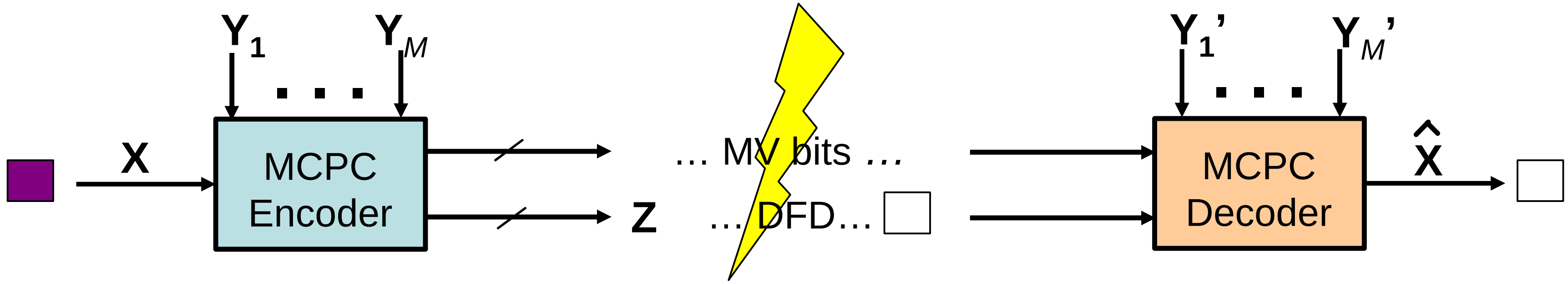
- Low-complexity motion-free encoder

- Need mechanism to detect decoding failure
- In theory: joint typicality (statistical consistency)
- In practice: Use CRC

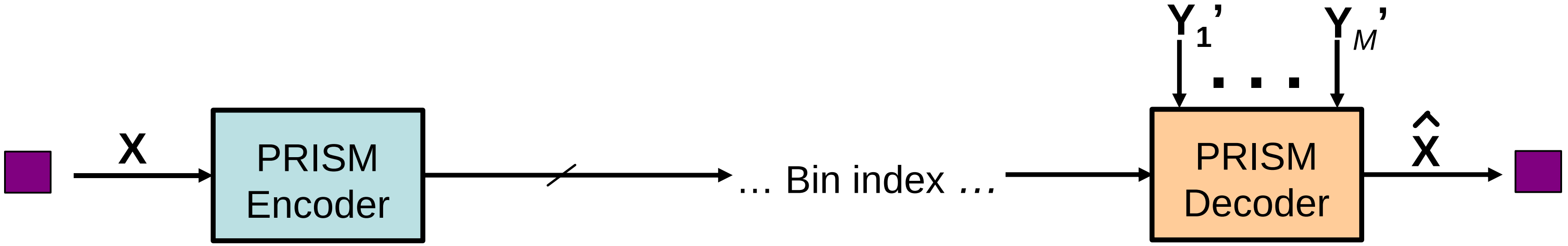
Need concept of "motion compensation at decoder"



Noisy channel: drift correction



MCPC: Channel errors lead to prediction mismatch and drift.



PRISM: Drift stopped if syndrome code (number of cosets) is “strong enough”

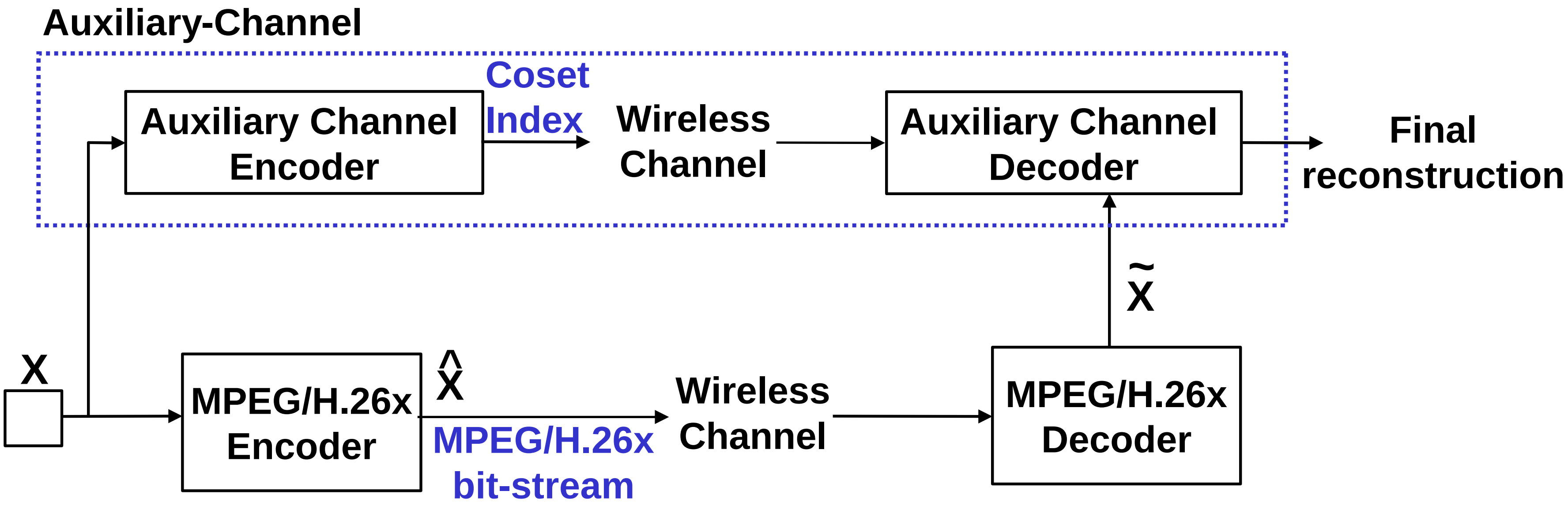
- All that matters: Targeted syndrome code noise

$\geq$ Correlation Noise + Induced Channel Noise + Quantization Noise

Enhancing the Robustness of MPEG Video

- What should be sent over this “auxiliary-channel” to get the best reconstructed video quality?
- Standard approaches:
  - Automatic Resend Request (ARQ):
    - Needs a feedback channel and has unbounded delay.
    - Does not scale in multicast setting.
  - Forward Error Correction codes (FEC):
    - Large latency required for good codes.
    - Good for error “prevention” but bad for drift correction.
- Proposed solution can correct drift with low latency.

Auxiliary-Channel Setup



- A secondary description of the video is sent over an “auxiliary-channel”.
- $\tilde{X}$ is a noisy version of X and will be used to help decode the auxiliary stream.

Simulation Results

- Qualcomm's channel simulator for CDMA 2000 1X wireless networks used.
- Predictive coder: H.263+ (obtained from UBC) & H.264.
- Auxiliary Channel rate fixed at 20% of total rate (3GPP recommendation).
- 3 systems:
  - PRISM.
  - Auxiliary channel
    - Full rate to FEC (RS code with 1 frame latency)
    - PRISM + FEC (RS code with 1 frame latency)

- Football
  (SIF, 1.7 Mbps, 15 fps, 8% error)
- Stefan
  (SIF, 2.2 Mbps, 15 fps, 5% error)

Aux Chan.
vs.
H.263+ FEC

PRISM
vs.
H.263+ FEC

PRISM
vs.
H.264

PRISM
vs.
H.264 FEC

Conclusions

- DSC-based video coding features:
  - Enhanced robustness
  - Flexible distribution of complexity (i.e. light encoding)
- Other relevant applications:
  - Scalable video coding (lack of state explosion at the encoder while avoiding drift)
  - Multicamera coding (without communication among cameras)

Thank you!

Lago di Como - Italy